

Möllers, Judith; Fritsch, Jana; Buchenrieder, Gertrud

**Article — Published Version**

## Farm and non-farm incomes of rural households in Slovenia: Canonical correlation analysis

South East European Journal of Economics and Business

**Provided in Cooperation with:**

Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale)

*Suggested Citation:* Möllers, Judith; Fritsch, Jana; Buchenrieder, Gertrud (2008) : Farm and non-farm incomes of rural households in Slovenia: Canonical correlation analysis, South East European Journal of Economics and Business, ISSN 2233-1999, University of Sarajevo, Faculty of Economics, Sarajevo, Vol. 3, Iss. 2, pp. 39-48, <http://dx.doi.org/10.2478/v10033-008-0013-9>

This Version is available at:

<http://hdl.handle.net/10419/234593>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

# Farm and Non-farm Incomes of Rural Households in Slovenia Canonical Correlation Analysis

Judith Möllers, Jana Fritzsich, Gertrud Buchenrieder\*

## Abstract:

**This paper provides an analysis of socio-economic characteristics and their influence on farm and non-farm incomes of rural households in Slovenia. With the canonical correlation analysis we use a methodological approach that offers a true multivariate procedure for both sides of the equation. It thus goes beyond a simple pair-wise correlation analysis and also beyond multiple correlation analysis. This rather rarely used statistical method offers interesting insights into many fields of analytical applications. Our results confirm that rural households usually turn towards non-farm employment if distress-push factors prevail. Besides insufficient farm incomes, large household sizes push households into non-farm diversification. Employment opportunities in the non-farm sector- regardless of whether triggered by distress in the household or demand in the non-farm sector- depend strongly on education.**

Keywords: Canonical Correlation Analysis, Slovenia, Rural Non-farm Sector, Diversification

JEL: D10, J20, R20

DOI: 10.2478/v10033-008-0013-9

## 1. Introduction

In rural areas of transition economies only the lucky ones can rely on agricultural incomes alone. Small-scale farming systems and unfavourable economic conditions lead to a need to take up additional non-farm employment to sustain the households' livelihoods. The transition process came along with a dramatic decrease in production and high unemployment rates. While a few stood to benefit from these economic changes, particularly in rural areas, many face unemployment, loss of life-time savings and consequently have to now endure poverty. Thus, the start of non-farm employment often is the result of 'distress-push' dynamics, itself often related to downward pressures on incomes from farming (Benjamin and Kimhi 2006).

The share of non-farm incomes in the income portfolio of rural households is known to be substantial, but is usually neither quantitatively investigated nor recorded in statistics of most of Central and Eastern European countries (CEEC). Greif (1997) estimates the contribution of non-farm incomes in CEEC as between 15 and 60% of all incomes. According to national statistics in Slovenia, almost one quarter of the workforce of Slovenian family farms is employed off-farm and as many as 72% refer

to their farm work as an additional or casual activity (SORS 2000). In Slovenia around 10% are employed in agriculture, but the share of gross domestic product (GDP) is lower, at around 3%. Yet almost 90% of the area is rural according to the OECD-definition, and almost 60% of Slovenia's population lives in these areas (Juvančič and Erjavec 2001). On average, family farms in Slovenia

### \*Möllers:

Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-str.2, 06120 Halle (Saale), Germany  
e-mail: moellers@iamo.de

### \*Fritzsich

Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-str.2, 06120 Halle (Saale), Germany

### \* Buchenrieder

Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-str.2, 06120 Halle (Saale), Germany

are of less than six hectares, with more than half of this area being grassland.

The continued high involvement of the rural population in agriculture, the structural imbalances caused by the transition process and significant regional disparities have increased interest from both policy makers and researchers in the rural non-farm economy. Knowledge of the driving forces of income diversification and factors influencing the access to farm and non-farm incomes is thus essential to better exploit its potential to overcome structural problems and to alleviate poverty and income disparity.

This paper provides an analysis of socio-economic characteristics and their influence on farm and non-farm incomes of rural households in two research regions in rural Slovenia. We analyse income portfolios and investigate how farm and non-farm incomes depend on household characteristics. With canonical correlation analysis we use a methodological approach that offers a true multivariate approach for both sides of the equation. It allows for the identification of significant relationships between two sets of variables and goes beyond a simple pair-wise correlation analysis, where the relationship between two variables is tested, and also beyond multiple correlation analysis, where the relation between one variable and a set of variables is analysed.

In Section 2 the analytical tool known as canonical correlation analysis is introduced. Section 3 presents results based on the empirical database from Slovenia.

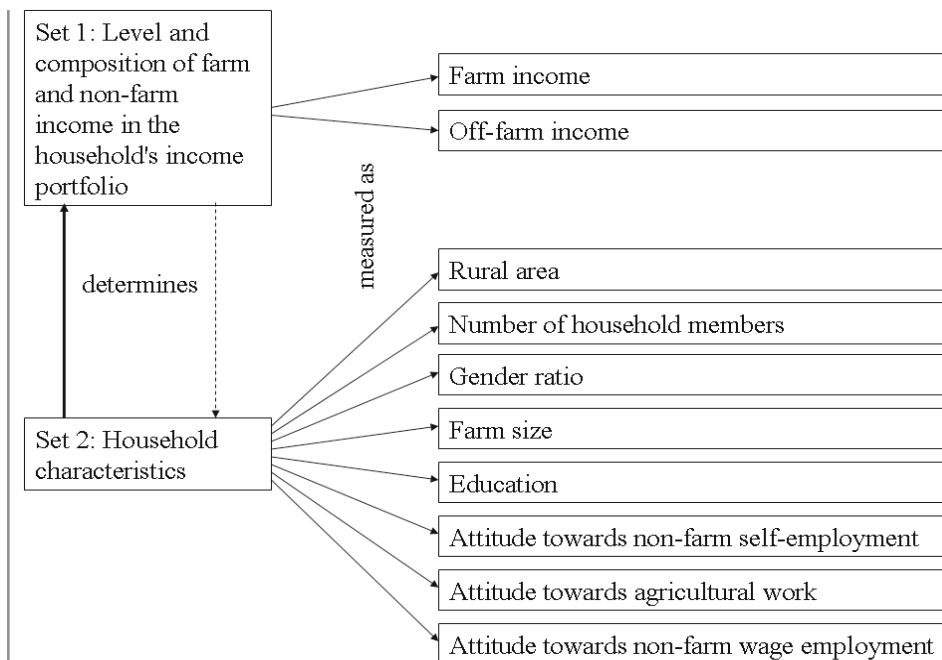
First, farm and income structures are briefly discussed and the variable sets for the correlation analyses are introduced. Following this is a presentation of the results of a standard and a canonical correlation analysis. The paper concludes in Section 4, also pointing out policy recommendations and remarks on the methodological approach.

## 2 Methodology

Canonical correlation analysis is used to reveal expected multivariate relations between variables that cannot be captured by the more commonly used methods of pair-wise correlation or multiple approaches such as logistic regression. We apply this method to test our main hypothesis, namely that the composition and level of rural households' income is to a large extent determined by their specific socio-economic characteristics (Figure 1). There may also be a reverse impact of households' income on their characteristics (dashed arrow in Figure 1), which will not be discussed here.

### 2.1 Database and Research Area

The empirical data against which we will test our hypothesis comes from a survey that was conducted in 2001 covering 120 rural Slovenian households<sup>1</sup>. Within the stratified sample, the households were selected



**Figure 1: Canonical correlation model to be tested in the analysis**

Source : Own figure

<sup>1</sup> We gratefully acknowledge the support of the European Commission for the funding of the Phare ACE project No. P98-1090-R.

according to their location and their income structure. Pomurska was selected as a typical rural region in Slovenia. Gorenjska represents a more peri-urban environment. Furthermore, the households are located in agriculturally more or less-favoured areas and four household types can be distinguished. These are (1) full-time farm households, (2) part-time farm households with at least one non-farm self-employment activity, (3) part-time farm households with at least one non-farm wage employment, and (4) households that have abandoned their farming activities. Within these groups the households were randomly selected.

Gorenjska lies in the north-west of Slovenia and is well connected to the capital Ljubljana. It is characterised by industrial production including steel, textiles and shoes. The unemployment rate is relatively low; at the time of the survey it was around 11% and below the national average. Pomurska, in the very north-east of the country, is little-industrialised. The non-farm labour market is based on the textile and food industries, but the textile industry suffered particularly from the transitional breakdown. The unemployment rate was almost 20% compared to about 14% in Slovenia<sup>2</sup>, and the number of people working on family farms in this region is, at 8.5%, clearly above average (2.9%) (SORS 2002; Erjavec et al. 2002a)

## 2.2 Canonical Correlation as Analytical Tool

As the procedure of the canonical correlation analysis is rather unknown and not often used in agricultural economics, we first briefly discuss this statistical method here. Hotelling (1935) was the first to describe the mathematics of canonical correlation. Good introductions into this procedure give Marinell (1990), Thompson (1984), Hair et al. (1998), and Stevens (2002). The main objective of the canonical correlation procedure is to identify significant relationships between two sets of variables. Each of these sets consists of at least two variables. This not only goes beyond a simple pair-wise correlation analysis where the relationship between two variables is tested, but also beyond multiple correlation analysis where the relation between one variable and a set of variables is analysed. Canonical correlation analysis provides a true multivariate approach for both sides of the equation. Its main advantages are thus that (1) a large amount of information on correlations that is otherwise only available pair-wise can be treated simultaneously and (2) the interpretation refers to only two sets of variables and is thus more condensed and intuitive.

For a canonical correlation analysis, the original set of  $t$  variables is divided into two sets, one with  $p$  and the other with  $q$  variables ( $t=p+q$ ). The canonical correlation

analysis looks for a linear combination (Equation 1 and 2) of the variables in each set so that the Pearson correlation between these linear combinations (canonical correlation  $Rc_m$ ) is maximized. In Equation 1 and 2,  $a_i$  ( $i=1 \dots p$ ) and  $b_j$  ( $j=1 \dots q$ ) stand for the coefficients in the linear combinations, and  $y_i$  and  $x_j$  represent the original variables.

Equation for set 1:

$$u_m = \sum_{i=1}^p a_i * y_i \quad (1)$$

Equation for set 2:

$$v_m = \sum_{j=1}^q b_j * x_j \quad (2)$$

The values of the linear combinations replace the original values in the calculation. A maximum of  $k$  pairs of linear combinations can be calculated with  $k=\min(p,q)$ . Each pair of  $u_m$  and  $v_m$  ( $m=1 \dots k$ ) is called the  $m^{\text{th}}$  canonical function; the linear combinations  $u_m$  and  $v_m$  themselves are the  $m^{\text{th}}$  canonical variates.

The interpretation of the results is done in four steps. First, the significance level of the canonical correlation coefficients ( $Rc_m$ ) is evaluated. As test statistic - Wilk's lambda - with a significance level of 0.05 is used. However, not only the significance level but also the magnitude of the canonical correlation coefficient should be considered in the context of the sample size. This is done in the second step of interpretation. Hair et al. (1998) recommend an absolute value of 0.5 as the lowest limit for an important correlation for a sample size of 120. Third, the canonical loadings are interpreted. Canonical loadings show the correlation between the original variable and its canonical variates. For canonical loadings the same limits are applied as for canonical correlations (0.5 for  $N=120$ ). Finally, the redundancy of the whole canonical analysis is assessed by the redundancy index because, as Stewart and Love (1968: 160) point out, it "is important to note that a relatively strong canonical correlation may be obtained between two linear functions, even though these functions may not extract significant portions of variance from their respective batteries." They define the redundancy index as:

Set 1:

$$RIv = \sum_{m=1}^k Rc_m^2 * Vv_m \quad (3)$$

<sup>2</sup>The unemployment rates refer to the total number of unemployed persons, both registered and those who consider themselves unemployed.

Set 2:

$$RIu = \sum_{m=1}^k R_{c_m}^2 * V_{u_m} \quad (4)$$

Whereby  $RIu$  and  $RIv$  represent the redundancy index for set 1 and set 2.  $R_{c_m}^2$  is the squared  $m^{th}$  canonical correlation.  $V_{u_m}$  and  $V_{v_m}$  relate to the proportion of the variance of set 1 or set 2 extracted by its canonical variate  $u_m$  or  $v_m$ .

The redundancy index measures the proportion of variance of set 1 predictable from set 2 ( $RIu$ ) and vice versa ( $RIv$ ). In most cases, only  $RIu$  is considered, since set 1 is defined as the set of dependent variables and researchers are more interested in the variance extracted from this set. There are no limits for an acceptable redundancy index; all results should, however, be assessed in the theoretical context of the research work and the empirical reality of the data set (Hair et al. 1998: 452).

To test the hypothesis, two sets of variables are defined. In our analysis the first shows the composition and level of farm and non-farm income in the households' income portfolios and the second describes household characteristics that are potentially influential with regard to these incomes. Set 1 comprises a variable representing income from agricultural activities, estimated on the basis of revenues and cost (see Erjavec et al. 2002), and non-farm income consisting of income from self-employment and waged employment. The income variables are measured in US dollar purchasing power parities (US\$ PPP). Set 2 includes eight variables that are defined as follows: (1) Dummy for location (0=household is located in the peri-urban region of Gorenjska, 1=household is located in the rural region of Pomurska), (2) number of all household members, (3) gender ratio (ratio of women between 16 and 64 years to all household members in the same age group), (4) farm size in hectares, (5) educational level of the main economic active person (MEA) ranging on a nine-point scale from 0=cannot read or write to 8=PhD studies<sup>3</sup> and (6) to (8) MEA's attitudes towards non-farm self-employment, agricultural activity and non-farm wage employment measured on a four-point rating scale (-1: negative attitude, 0: indifferent, 1: positive with slight reservation, 2: very positive attitude).

The data evaluation starts with descriptive statistics. In a second step, correlation coefficients (Pearson's correlation coefficient) for each pair of variables are calculated and tested for significance. Finally, canonical correlation analysis as a multivariate approach is adopted to test for significant relations between the composition of farm and non-farm income in the household's income portfolio (set 1) and the above described household characteristics (set 2). All calculations were done with the SPSS software package.

<sup>3</sup>Nevertheless, the highest identified educational degree was a Bachelor.

## 3 Empirical Results

### 3.1 Descriptive Statistics

Household characteristics are assumed to directly impact income strategies. Therefore a short overview of the variable sets used and some additional information on incomes, farm sizes, education, attitudes and perceived access restrictions to labour markets will be provided. Table 1 shows descriptive statistics for all variables in the analysis.

The average yearly per-capita income in the Slovenian sample was found to be 3,900 € (or 6,500 US\$ PPP), which is about 60% less compared to the national average (EBRD 2002). This significant difference can be explained, on the one hand, by the generally higher poverty risk in rural areas and, on the other hand, by the inclusion of Pomurska, the poorest region in the country in the sample. Non-farm incomes play an important role for rural households in Slovenia. On average they contribute over 40% to the total income portfolio; Figure 2 depicts regional income portfolios as found in the empirical data set. The most important branches of the rural non-farm sector are trade and rural tourism followed by metallurgy, the food industry and tailoring.

Farms are, compared to the national average, well endowed with land; the farm size is on average 13 ha. The share of farming income is around 30%, with the remainder made up of pensions, social payments, interest payments, etc. Interestingly, farm incomes are more important for the better-off income groups, pointing at predominantly distress-push motivations for those who take up non-farm employment. The access to better-paid non-farm occupations is closely related to education level. Bojnec et al. (2003), for example, show that those who are able to avail themselves of the rural non-farm labour market are characterised by a higher education. The observed re-allocation of labour back into the agricultural sector, however, is, related to high unemployment in the rural non-farm economy and also to old age. The agricultural sector in Slovenia thus acts as a buffer (Bojnec et al. 2003). In general, the educational level in rural Slovenia is relatively high. More than 60% of the sample households have at least one household member who graduated from a secondary school; 30% of all active household members have gone through a professional education.

Employment decisions are furthermore influenced by personal attitudes and by perceived constraints, for example, with regard to the access to the local labour market. The data reveals that the Slovenian rural population is very much inclined to non-farm employment, particularly to wage employment. Seventy-six percent of all respondents have a very positive attitude towards non-farm employment. At the same

time, Slovenians appreciate farming and the rural lifestyle. Less than 8% of the respondents gave negative ratings, but the analysis of the age groups shows that particularly the young generation in the age group from 16 to 25 has fewer positive attitudes towards agriculture. The most important constraints as perceived by the households, hindering them from taking up non-farm employment, are – besides a lack of time and labour capacity – lacking equity capital, difficulty in obtaining credits, unfavourable policies and administrative barriers as well as lacking infrastructural and market connections. The Kolmogorov-Smirnov test for normal distribution shows that, except for the number of household members, all other variables are not normally distributed (Table 1).

	Min.	Max.	Median	Test statistic <sup>1)</sup>	Sig. <sup>2)</sup>
<b>Set 1</b>					
Farm income (US\$ PPP)	-13,174	69,900	3,859	2.10	0.000
Non-farm income (US\$ PPP)	0	49,773	9,865	1.48	0.026
<b>Set 2</b>					
Rural area (1=rural, 0=else)	0	1	0.5	3.73	0.000
Number of household members	1	9	5	1.24	0.091
Farm size (total ha)	0.0	67.0	9.1	1.92	0.001
Household gender ratio	0.0	1.0	0.5	2.58	0.000
Education (nine-point scale) <sup>3)</sup>	1	6	3	2.38	0.000
<b>Attitude towards... (four-point scale)<sup>4)</sup></b>					
Non-farm self-employment	-1	2	1	3.48	0.000
Agricultural activities	-1	2	1	3.48	0.000
Non-farm wage employment	-1	2	1	5.15	0.000

Source: Own calculations with data from EC Phare ACE project No. P98-1090-R.

Note: <sup>1)</sup> Kolmogorov-Smirnov Z

<sup>2)</sup> A significance level smaller than  $p < 0.05$  indicates that the data are not normal distributed.

1 US\$ PPP = 133.8 Slovenian Tolar in 2001 (World Bank 2003).

30 households display negative farm incomes that are

compensated through non-farm incomes or unearned incomes, e.g. old-age pensions.

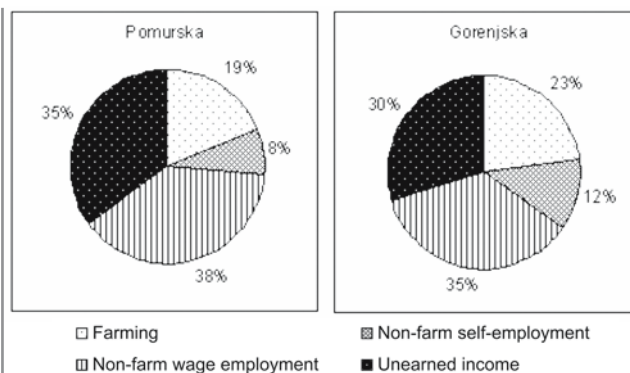
Set 1: Depicts the average farm and non-farm income in 2001.

Set 2: Depicts selected household characteristics with a potential influence on the income variable set; the first and the last two categories of the education variable are empty, which means no MEA has the respective educational level.

<sup>3)</sup> 0=no studies and cannot read or write, 1=no studies but can read or write, 2=elementary school, 3=vocational school, 4=secondary school, gymnasium, 5=college, 6=graduate studies (university B.Sc.), 7=M.Sc. studies (university), 8=Ph.D. studies (university). The categories 0, 7, and 8 were not mentioned from the respondents.

<sup>4)</sup> -1 = negative attitude; 0=indifferent attitude, 1=positive attitude, 2=very warm feeling about the respective activity.

**Table 1: Overview statistics for the Slovenian rural households, 2001**



Source: Data from EC Phare ACE project No. P98-1090-R.

Note: The averages are calculated on the basis of income portfolios of individual households. Unearned incomes = pensions, social security payments, child benefits, etc.

**Figure 2 Regional income portfolios**

### 3.2 Correlation Analysis

The correlations between all pairs of variables were calculated using Pearson's correlation coefficient and tested for significance at the 0.05 level. This analysis provides a priori information about relations between the composition and level of household's income on the one side and household characteristics on the other side for the canonical correlation analysis that will be employed in Section 3.3. In Table 2, all correlation coefficients are depicted and significant ones are marked with asterisks. All variables show at least one significant correlation with another variable. As all variables appear to yield some information, we will include all in the canonical correlation analysis. Seventeen of the 45 pair-wise correlations are significant. However, the absolute values of the coefficients

indicate that no strong relations exist. We will discuss the significant relations from the pair-wise correlations in the following section in more detail.

### Correlations within Set 1

Farm income is negatively correlated with non-farm income. This indicates that people who earn a satisfying level of income from farming participate less in non-farm work. People who have smaller farms tend to embark more upon non-farm work.

### Correlations within Set 2

The correlation between the rural area variable and the number of household members is negative. This indicates that households in the rural area of Pomurska have fewer members than those in the peri-urban area of Gorenjska. The Mann-Whitney test confirms that this correlation is significant (Mann-Whitney  $U=1421.0$ , significance level=0.044). The rural location of the household and the educational level of the MEA are negatively correlated. Although Slovenia is a relatively small country with good access to urban centres, the educational situation is better in Gorenjska, which is less remote and well connected to urban centres.

The number of household members is positively correlated with the gender ratio, the farm size and the MEA's attitude towards non-farm self-employment. Larger farms can provide an adequate livelihood for more family members. Smaller farms often can, if at all, sustain only the farm owner and the core family. More positive attitudes towards non-farm self-employment in larger households may be caused by the fact that self-employment is seen as an additional income source that fits in well with farm activities. Particularly in larger households it provides an opportunity to productively employ additional family members in the workforce. Moreover, the need for supporting more household members could also result in the decision to diversify income by self-employed activities. Farm size is positively correlated with the MEA's attitude towards agricultural activities. This demonstrates that owners of larger farms, who often obtain most of their incomes from farming, identify themselves with their profession more than small-scale farmers.

An interesting correlation arises between education and attitudes towards agricultural work. The negative correlation coefficient indicates that better educated people consider farm work less desirable. At the same time, the correlation between the MEA's attitudes towards non-farm wage employment and his educational level is positive. Investments in higher education frequently render non-farm waged employment more attractive.

### Correlations between variables of set 1 and set 2

As expected, farm size is positively correlated to farm income. Moreover, higher farm income positively interacts with the attitude towards farming. This indicates that people who earn sufficient agricultural income enjoy their farm occupation more than others. Non-farm income is positively correlated with the number of household members, education and attitudes towards non-farm waged employment. Larger households tend to diversify more into non-farm activities. On the one hand, they are obliged to do so because they must support more persons. On the other hand, they possess the manpower to embark upon additional activities. Access to non-farm work, particularly better-paid jobs, depends strongly on education.

On average, the biggest share of income comes from non-farm work in the sample (Figure 2). A positive correlation of the educational level exists with non-farm income but not with farm-income. Higher education may give access to better-paid and more attractive jobs. A positive attitude towards non-farm waged employment goes along with significantly higher non-farm income, whereas the attitude towards non-farm self-employment shows no significant influence. Qualitative results show that the rural population in Slovenia is particularly fond of wage employment as it is less risky and comes along with social security schemes (Möllers 2006). Accordingly, the majority of non-farm income comes from waged employment (63% of all households have such an income). Only 23% of the households are engaged in non-farm self-employed activities. Non-farm income is negatively correlated with the rural area dummy, farm size and attitudes towards agricultural work. The opportunities for non-farm income, especially from waged employment, are better in the peri-urban Gorenjska region.

Farm size and attitudes towards farm activities are negatively correlated to non-farm incomes. In households with higher non-farm incomes, the attitudes towards agricultural work are less positive. On the one hand, if agricultural work does not provide for a certain living standard, the attitudes towards farming could turn negative. On the other hand, when a household has to choose whether to concentrate on farming or to start non-farm employment, the decision for non-farm work is pushed forward by less positive attitudes towards farm activities.

In summary, the results of the pair-wise correlation analysis show that non-farm income is influenced by more variables than farm income. For the level of farm income, the farm size and attitudes towards farming activities are important, whereas non-farm activities are prompted by a large number of household characteristics. All correlations could be interpreted reasonably but, as mentioned above, it is difficult to come up with a clear picture of relations from pair-wise correlations. Therefore, a canonical correlation is calculated in the next step.

	Farm income	Non-farm income	Rural area	Number of household members	Gender ratio	Farm size	Education	Attitudes towards non-farm self-employment	Attitudes towards farm activities	Attitudes towards non-farm wage employment
Set 1										
Farm income	1.00									
Non-farm income	-0.33*	1.00								
Set 2										
Rural area	0.07	-0.22*	1.00							
Number of household members	0.15	0.23*	-0.21*	1.00						
Gender ratio	0.01	-0.13	0.02	0.25*	1.00					
Farm size	0.69*	-0.27*	0.12	0.23*	-0.04	1.00				
Education	-0.02	0.37*	-0.09*	0.04	0.00	0.02	1.00			
Attitude towards non-farm self-employment	0.12	0.14	-0.15	0.20*	-0.07	0.08	0.06	1.00		
Attitude towards farm activities	0.22*	-0.22*	0.17	-0.00	-0.07	0.26*	-0.26*	0.12	1.00	
Attitude towards non-farm wage-employment	-0.14	0.18*	-0.10	0.11	0.00	-0.08	0.28*	0.04	0.00	1.00

Source: Own calculations with data from EC Phare ACE project No. P98-1090-R.

Note: \* significant on the 0.05-level

**Table 2: Pearson's product-moment-correlation coefficient for all variables in the analysis**



### 3.3 Canonical Correlation Analysis

The results presented above show multiple relations between all variables. Therefore, there is no simple answer to the question of which household characteristics determine the level and composition of farm and non-farm income in the household's income portfolio. By using the statistical procedure of canonical correlation analysis we look at all correlations simultaneously and thus identify important household characteristics.

Set 1 involves two variables; hence two canonical correlation coefficients are calculated. Table 3 lists the canonical correlation coefficients and the canonical loadings. Both canonical correlations are highly significant and larger than 0.5. Thus, they are important and further interpretation is useful.

	1 <sup>st</sup> canonical function	2 <sup>nd</sup> canonical function
Set 1		
Farm income	-0.97	-0.24
Non-farm income	0.54	-0.84
Set 2		
Rural area	-0.20	0.40
Number of household members	-0.09	-0.63
Farm size	-0.96	-0.21
Gender ratio	-0.06	0.27
Education	0.12	-0.59
Attitude towards non-farm self-employment	-0.09	-0.42
Attitude towards agricultural activity	-0.35	0.20
Attitude towards non-farm wage-employment	0.22	-0.17
Canonical correlation	0.71	0.52

Source: Own calculations with data from EC Phare ACE project No. P98-1090-R.

Note: Bold figures indicate important canonical loadings. A loading is considered important when it reaches at least a value of 0.5 (Hair et al. 1998).

**Table 3: Canonical loadings**

Looking at the canonical loadings for set 1, it becomes obvious that the first canonical variate mostly represents farm income, whereas the second represents non-farm income. Non-farm income is also correlated with the first canonical variate. Higher non-farm incomes result in or are caused by lower farm incomes. In other words, the two income components substitute for each other.

The first canonical variate for set 2 is determined by farm size only. The negative signs indicate that larger farms provide higher farm incomes. Accordingly, households with less agricultural land are more eager to enter into non-farm activities and have higher absolute non-farm incomes. The second canonical variate for set 2 has high loadings for the number of household members and the educational level. Non-farm income rises with the number of people living in the household and with the educational level of the MEA.

As explained in Section 2, not only the canonical correlation coefficient and the loadings should be considered but also the redundancy index. The redundancy indices are 0.410 for set 1 and 0.115 for set 2, respectively (Table 4).

	1 <sup>st</sup> canonical function	2 <sup>nd</sup> canonical function
Set 1		
Farm income	<b>-0.97</b>	-0.24
Non-farm income	<b>0.54</b>	<b>-0.84</b>
Set 2		
Rural area	-0.20	0.40
Number of household members	-0.09	<b>-0.63</b>
Farm size	<b>-0.96</b>	-0.21
Gender ratio	-0.06	0.27
Education	0.12	<b>-0.59</b>
Attitude towards non-farm self-employment	-0.09	-0.42
Attitude towards agricultural activity	-0.35	0.20
Attitude towards non-farm wage-employment	0.22	-0.17
Canonical correlation	<b>0.71</b>	<b>0.52</b>

Source: Own calculations with data from EC Phare ACE project No. P98-1090-R.

**Table 4: Redundancy indices**

Set 1 was defined as the set of dependent variables that are explained by the variables of set 2 as independent variables, therefore only Rlu is important for interpretation. 41% of variance in set 1 is explained by all canonical variates of set 2. This leaves nearly 60% unexplained. However, in the context of the hypothesis and a wide variability in the original data, this result is satisfactory.

The network of multiple relations could be described by two canonical functions. In comparison to the correlation analysis, the canonical correlation analysis provides a much clearer and comprehensive picture of relationships. It results in three main findings:

1. Only three of the included household characteristics influence the composition and level of household's income, i.e. farm size, number of household members and the educational level of the MEA.
2. Two more variables, (1) the region and (2) attitudes towards self-employment show loadings near the lower limit of 0.5 and thus could have an impact. However, due to the small sample size, the variables were not further interpreted.
3. The gender ratio and attitudes towards agricultural and non-farm waged employment do not influence the composition and level of household's income significantly.

#### 4. Conclusions

Not surprisingly, household characteristics influence the composition and level of household income significantly. While the commonly used pair-wise correlation analysis gives some insights into possibly important dependent variables and their relationship, it fails to provide a clear picture of important determinants. The procedure of canonical correlation, however, reveals that out of the potential determinants only three influence the composition and level of household's income significantly. These are the farm size, the number of household members and the educational level of the main economic active household member.

We find that farm size and income from agricultural activities are positively correlated. Farm incomes are almost solely influenced by the size of the farm. Otherwise, non-farm incomes are negatively influenced by the farm size and thus larger farms have a lower probability that household members engage in the non-

farm sector. The two components of household income are related, i.e. higher non-farm incomes result in or are caused by lower farm incomes. In other words, the two income components substitute for each other.

Our results confirm that rural farm households usually turn towards non-farm employment if farm incomes are not sufficient to support their livelihoods. Another factor that pushes households into non-farm diversification is the size of the household: larger households have significantly higher non-farm income than smaller ones. However, we also find that the incomes that can be gained in the rural non-farm sector clearly depend on education. Higher educational levels open possibilities in the non-farm sector, whereas it does not show any impact on farm incomes. All other potential determinants of farm and non-farm income of rural households in Slovenia turned out to be insignificant when analysed simultaneously with the canonical correlation analysis.

From this we can deduct a couple of recommendations for policy makers: In countries with predominantly small-scale farms, rural development measures should address two different target groups. First, farms with a potential to develop into viable and competitive sizes should be in the focus of agricultural policies. It is important to provide incentives that facilitate factor market mobility and particularly allow these farms to grow; i.e. land transfers (sale or rental) from those farm families who leave the sector should be facilitated and promoted. Second, rural development policies that foster the local non-farm economy are needed. Often the uptake of non-farm jobs is mainly restricted by underdeveloped local job markets as well as a lack of appropriate general and professional education. Therefore it is necessary to provide incentives for business start-ups and to increase the attractiveness of a region for establishing industries. A further focus should be put on improving the employability of the labour force. In order to increase employment of those with a low level of education, or of those with knowledge and skills that are not in demand in the labour market, it is necessary to emphasise active forms of assistance, i.e. training and education in accordance with changing labour market needs.

The methodological approach of using the canonical correlation sheds light on complex relationships between potential determinants of rural income levels and income composition that are otherwise difficult to disentangle. Being a true multivariate approach for both sides of the equation, it has clear advantages compared to the more commonly used pair-wise correlations or multiple regressions. □

## References

- Benjamin, C. and A. Kimhi 2006. Farm work, off-farm work and hired farm labour: estimating a discrete-choice model of French farm couple's labour decisions. *European Review of Agricultural Economics* 33 (2): 149-171
- Bojnec, S., L. Dries and J. F. Swinnen 2003. Human Capital and Labour Flows out of the Agricultural Sector: Evidence from Slovenia. *Proceedings of the 25th International Conference of Agricultural Economists (IAAE)*, 16-22 August 2003, Durban.
- EBRD 2002. *Transition Report Update*. European Bank for Reconstruction and Development, London: EBRD.
- Erjavec, E., L. Juvančič, O. Oblak, M. Kožar, B. Moljk, P. Kolarek and S. Kavčič 2002. *Economic Diversification and Labour Adjustment in the Rural Households in Slovenia*. Working Paper EC Phare Ace Research Project P98-1090-R.
- Greif, F. 1997. *Off-farm income sources and uses in transition economies*. Federal Institute of Agricultural Economics, Vienna und FAO/REU, Rome.
- Hair, J. F., W.C. Black, B.J. Babin, R.E. Anderson and R.L. Tatham 1998. *Multivariate Data Analysis*, 5th edition, Prentice Hall International, Upper Saddle River, New Jersey.
- Hotelling, H. (1935). *The Most Predictable Criterion*. *Journal of Educational Psychology*. Vol. 26: 139-142.
- Juvančič, L. and E. Erjavec 2001. *Rural development in Slovenia – A brief outline*. Working Paper EC Phare Ace Research Project P98-1090-R.
- Marinell, G. (1990). *Multivariate Verfahren*, 3rd edition. Oldenburg Publisher, Munich.
- Möllers, J. 2006. *Außerlandwirtschaftliche Diversifikation im Transformationsprozess. Diversifikationsentscheidungen und -strategien ländlicher Haushalte in Slowenien und Mazedonien*. *Studies on the Agricultural and Food Sector in Central and Eastern Europe*, Vol. 35, Halle (Saale).
- Stevens, J. P. 2002. *Applied Multivariate Statistics for the Social Science*. 4th edition, Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Stewart, D. and W. Love 1968. *A General Canonical Correlation Index*. *Psychological Bulletin*. Vol. 70, 160-163.
- SORS 2000. *Agricultural Census 2000*. Statistical Office of the Republic of Slovenia Ljubljana: SORS.
- SORS 2002. *Statistical Yearbook of the Republic of Slovenia 2002*. Statistical Office of the Republic of Slovenia, Ljubljana: SORS.
- Thompson, B. 1984. *Canonical Correlation Analysis*, Sage University Paper series on Qualitative Applications in the Social Sciences No. 47, Sage Publications, London.
- World Bank 2003. *World Development Indicators*. Washington, D.C.: The World Bank.